

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 21:32:49 : Search time 3319.91 Seconds
(without alignments)
11301.897 Million cell updates/sec

Title: US-09-719-017a-2

Perfect score: 1793

Sequence: 1 gattcccccgttgacacatla.....caattactcaatgcgcgcg 1793

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_others:*

33: em_htgc_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
------------	-------	-------	-------	--------	----	----	-------------

1	1793	100.0	1793	6	AX008753	AX008753 Sequence
2	1793	100.0	1793	6	AX025996	AX025996 Sequence
3	1132.8	63.2	1727	1	AFMANL	D13419 A. faecalis
4	1097.6	61.2	1727	6	ECORGNB	E07629 gDNA encodi
5	500	27.9	7508	1	ECORGNB	J01695 E.coli rRNA
6	500	27.9	7508	1	ECORGNB	V00348 E. coli rib
7	500	27.9	11679	1	AE000471	AE000471 Escherich
8	500	27.9	11918	12	CUV2488	U72488 Cloning vec
9	500	27.9	176195	1	ECOM89	U00006 E. coli chr
10	421	23.5	860	12	SYNRGNABP	K00764 e.coli rna
11	404	22.5	5094	12	XXU13859	U13859 pK232-8-cl
12	404	22.5	6733	12	AF216802	AF216802 Shuttle v
13	404	22.5	6984	12	AF216803	AF216803 Shuttle v
14	404	22.5	8565	12	SYNPT61T	M29896 Cloning vec
15	404	22.5	9113	12	CVPCG1408	U21228 Promoter-pr
16	404	22.5	13331	12	AF209190	AF209190 Shuttle v
17	403.6	22.5	502	23	E09023	E09023 3'-terminal
18	401.6	22.4	7965	12	AF327719	AF327719 Cloning v
19	401.6	22.4	11299	12	AF327720	AF327720 Cloning v
20	396.4	22.1	2491	1	ECOBTRA	M10123 E.coli b1ra
21	388.4	21.7	4204	6	E02679	E02679 All sequenc
22	388.4	21.7	4207	6	E02104	E02104 DNA sequenc
23	388.4	21.7	4222	6	E02098	E02098 DNA sequenc
24	388.4	21.7	4466	6	E02001	E02001 DNA sequenc
25	388.4	21.7	4466	6	E06722	E06722 Plasmid pTP
26	369	20.6	7416	6	AX354080	AX354080 Sequence
27	336	18.7	5641	6	AX113748	AX113748 Sequence
28	336	18.7	5670	6	AX113749	AX113749 Sequence
29	336	18.7	5826	6	AX113746	AX113746 Sequence
30	336	18.7	6071	6	AX113747	AX113747 Sequence
31	332.8	18.6	4176	12	AF028838	AF028838 Cloning v
32	332.8	18.6	4176	6	A13038	A13038 Artificial
33	332.8	18.6	4176	12	SYNTRC99A	M22744 Cloning vec
34	332.8	18.6	4176	12	XXU13872	U13872 pTic99a c10
35	332.8	18.6	4214	6	AX021488	AX021488 Sequence
36	332.8	18.6	4239	6	A29289	A29289 pSEC-Bpl co
37	332.8	18.6	4357	6	AX072889	AX072889 Sequence
38	332.8	18.6	4476	6	AR102989	AR102989 Sequence
39	332.8	18.6	4476	6	AR172783	AR172783 Sequence
40	332.8	18.6	4566	12	AF177933	AF177933 Cloning v
41	332.8	18.6	4593	6	AR102988	AR102988 Sequence
42	332.8	18.6	4593	6	AR172782	AR172782 Sequence
43	332.8	18.6	4593	12	CVPRK2332	X70478 Plasmid PK
44	332.8	18.6	4593	12	XXU02439	U02439 Cloning vec
45	332.8	18.6	4635	12	AF177932	AF177932 Cloning v

ALIGNMENTS

RESULT 1

AX008753 LOCUS AX008753 1793 bp DNA linear PAT 06-SEP-2000

DEFINITION Sequence 2 from Patent WO9964607.

ACCESSION AX008753

VERSION AX008753.1 GI:9996244

KEYWORDS

SOURCE

ORGANISM

synthetic construct.

synthetic construct

artificial sequence.

REFERENCE 1 (bases 1 to 1793)

FAVRE-BULLE O., GUILTON C. and PIERRARD J.

TITLE Industrial method for producing heterologous proteins in e.coli and strains useful for said method

JOURNAL Patent: WO 9964607-A 2 16-DEC-1999;

FAVRE-BULLE OLIVIER (FR); GUILTON CAROLE (FR); PIERRARD JEROME (FR); RHONE-POULENC NUTRITION ANIMAL (FR)

FEATURES

source

1..1793

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="cassette d'expression"

123..1193

CDS

```
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC07435.1"
/db_xref="GI:3996245"
/translation="MQTRKIVRAAAVQASPNYDLATGVKDTIELARQARDEGDLIV
FGEFMPLPGYPRHVMIGAPMSLKYSAARYANSLSJDSSEFRIQAAATGILFIALGY
SERGGSLYLQCLIDDKQMLMSRRLKPHVEVTVGEVARDLIYSPDELGRVGA
LCQNEHLSPLSKVLYLSQHEAIHTAAMSFSLYSQADALSAKYNMAASQIYSEVGC
FTTAASTVYQDETLDMLEVBHNASILKVGSGSSMIFAPDQRTTAPYLPDPAEGILTA
DLNMEETAFKAIINDPVGHYSKPEATRIYLDLGRHREPMTRVHSSVIDEAPRPHVQS
TAAAPVAVSQTDSDTLVQEPS"
BASE COUNT      412 a      527 c      478 g      376 t
ORIGIN
Query Match      100.0%; Score 1793; DB 6; Length 1793;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 gaattccctgttgacaatcaatcatcgaactagtaactagtaacgacgttgctgcag 60
Db      1 GAATTCCCTGTTGACAAATTAATCATCGAAGCTAGTAAGTACGACGCTTGCTGCAG 60
QY      61 tcgacctgcagccaagcttggcatalacatcaatgttatacbaaggaatacttaac 120
Db      61 TCGACCTCGAGCCAAAGCTTGGGCATACATTCATCAATGATTGTTATGAAGAAATACTTAC 120
QY      121 atatgcagacaagaanaaatcgtccggcgacgcgcgttaagagccgcctctcccaactaag 180
Db      121 ATATGCAGACAAGAAANAATCGTCGGGCGACGCGCGGTACAGGCGCGCTCTCCCACTACG 180
QY      181 atctgcagcaaggtgttgaataaacatltgactgtcgtcagagcccgcgatgaaggtc 240
Db      181 ACTCTGGCAAGGGGTGTTGATTAACCATGTGAGCTGGCTGTGAGGCGCCGCGATGAGGGCT 240
QY      241 gtgacctgactglttltgttgaacacctggtccggcgatacccttccacgtctggtcgtg 300
Db      241 GTGACCTGACTGTTTGTGTGTAACCTGGCTGCCGGCATTCCTTCCACGCTCTGGCTGG 300
QY      301 ggcacacggcctgtgtcgtcgtgaataacgtgtccgcgtactactgcaactcgtctcgtcgtg 360
Db      301 GCGCACCGGCTGTGCTGCTGGAATACAGTGCCTGCTACTATGCTCAACTGCTCTGCTGG 360
QY      361 acagtgcaagagtltcaacgcacatltgccagccgcagcagacacttgggtatattcatcgcac 420
Db      361 ACAGTGCAGAGTTCACACGCAATTCGCCACAGCGCGACGACCTTGGGTATTTTCATCGCAC 420
QY      421 ttggttatagcgaagcgcagcgcgcgcgttaccttggggccaatgtgacctgacgcagacaa 480
Db      421 TTGGTTATATAGCGAGCGCACGCGCGGCGAGCTTACCCTGGGCCAATGCCCTGATCGACGACA 480
QY      481 agggcacaagtgtcgtgtgtcgtcgcgaacactcaaacctcacatgtltgagcgacgcgtgt 540
Db      481 AGGGCACAAGTGTGTGTGTGCTGCCGCAAACTCAAACTACATGTGTGAGCGCACCGCTGT 540
QY      541 ttgtgtgaaggltatgcgcccgagatcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
Db      541 TTGTGTGAAGGTTATGCCCGAGATGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY      601 cccctgtgcgtcgtggagagacacctgtccctctgagcaagaagcgcgtgtatctcccgcaag 660
Db      601 CCTTGTGCTGCTGTGGAGACCTGTCCCTTGTGAGCAAGTACGCGCTGTCTCCACGACAG 660
QY      661 aagcactacatgtgcgcgtcgtgcgcgtcttcttgcgtgtgacgcgcaaacggcccatgtcgc 720
Db      661 AAGCAATTCACATTGTGCGCGCTGCGCGCTCTTTTCCGTGTACAGCGACAGCGCCCATGCGCC 720
QY      721 tcaagcgcaaggltgaacatgtgtcgtcgcgaatcataltcggtltgaagcgcaagltgcttca 780
Db      721 TCAGGCGCCAAAGGTGAACATGTGCTGCCCTCGCAATCATTCGGTGTGAAGCGCAGTGCTTTA 780
QY      781 ccatgcgcgcgcagcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
```

```
|||||
Db      781 CCATGCGCGCCAGCAGTGTGTCACCCAGAGACACTGCAATGCTGGAAGTAGTGAAC 840
QY      841 acaagcctccctgtctgaagatgtggcgcgcgcgtgttcatgtatlttttgcgcgcgcgcgcgc 900
Db      841 ACAAGCCTCCCTGCTGTGAAGATGGGCGGCGGAGTGTCCATGTATTTTTCGCGCGGAGAC 900
QY      901 gacactgtctccctcactgcacacacagatgtccgaagcctgtatcattgcgaatccgaaca 960
Db      901 GCACATTGGCTCTCCACTGCGCACACAGATGCCGAAGGCTGATTCATTGCCATCTGGAACA 960
QY      961 tggaaagaattgtccttcgcgaagcgatcaacgacctgtgtggccactactccaaacccg 1020
Db      961 TGGAAGAATTTGCTTTCGCGCAAGGCGATCAACGACCTGTGGGCGCATCTCCAAACCCG 1020
QY      1021 agggcaccgtctgttactgtgacgtggcgacacctgtgaagccatgaactcgggtgtatcca 1080
Db      1021 AGGCCACCGCTGTGTACTGTGACCTGGGGCACCGTGAAGCCCATGTGACGTGGGTGCAATTCCA 1080
QY      1081 aaagcgtatccaggaagaagctcccgagccgacgttcaaaqtaacgctgcgcgcgtcgtc 1140
Db      1081 AAAGCGTATCCAGGAAGAAGCTCCCGAGCGCGACGTCGAAMAATGAGGCTGGCGCGTGTG 1140
QY      1141 ccgtcaagccagactcaaggaactgtgaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1200
Db      1141 CCGTCAGCCAGACTCAGGACTCGGATGAGCTACTGTGCAAAACCGTCTGACCCCAAA 1200
QY      1201 agatgaagaagccggcggaactgtcgcgtgttacttccctctgtgcgcgcgcgcgcgcgcgc 1260
Db      1201 AGATGACAAAGGCCCGGCAAACTGTCCGGCTTATTTCTTCTGCTGTCGTCGCGGATCCACT 1260
QY      1261 agtctagaagtcgaactcgaagcaltgcaagcttgggtcccaacttgaaccatltgcgaact 1320
Db      1261 AGTTCTAGAGTGTGACCTGACGAGCTGCAAGCTTGGGTGCCACCTGACCCCATGCCGAAT 1320
QY      1321 cgaagatgaagccgcgtgtgacgcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1380
Db      1321 CAGAAGTGAAGCGCGGTACCGCGGATGTGTGTGGGTCTCCCATGCGAAGTAGGA 1380
QY      1381 actgcagacatcaataaagaagcgtcagtcgaagaagcttgggcttctgttctatc 1440
Db      1381 ACTGCCAGGATCAATTAAGAAAGCTCAGTGCAGAAAGACTGGGCTTGTGTTTATTC 1440
QY      1441 tgtgttctgtgcgtgtgaagccttccctgaagtaagacaatccgcgcgcgcgcgcgcgcgcgc 1500
Db      1441 TGTGTTTGTGCGTGAACGCTCTCTCTGAGTAGACAAATCCCGCGGAGCGGATTTGAAC 1500
QY      1501 gttgcgaacaaacggcccggaaggtgtggtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1560
Db      1501 GTTGCGAACAAAGGCCCGGAGGAGTGCGGAGGAGGCGCGCCCATTAACCTGCCAGGCAT 1560
QY      1561 caaatlaagcagaagccatccctgaacgatagtgccttltgtcgttcttcaaaacttccct 1620
Db      1561 CAATTAAGCAAGAAGGCATCTCTGAAGATGGCTTTTGGCTTCTCAAAACTCTTCTCT 1620
QY      1621 gtgcgtcatactacaagcatalcccccacagatacgtgtaaactagcctcgttctgtcatc 1680
Db      1621 GTGCTCATATTCTAACAGGCATCCCGCCACAGATPAGGTAAACTCTGCTGTTTTGTGATC 1680
QY      1681 aggaagaagcagctatgaacacactccttaaaccttgaagaacatttggcactgtatcat 1740
Db      1681 AGGAAGAAGCAGCTATGACCACTCTTAHAACCTTGGAACACATTGGCATTTATATATANT 1740
QY      1741 gctcagacatgtatgtgtcgaagaagacagaacaattactaatgtccgcgcgcgcgcgcgcgc 1793
Db      1741 GCTCAGCAATTTGATGTGTGCCGAGACGACGAACAACATTAATCAATGCCGCGG 1793
RESULT      2
AX025996      1793 bp      DNA      linear      PAT 16-SEP-2000
LOCUS      AX025996      1793 bp      DNA      linear      PAT 16-SEP-2000
DEFINITION      Sequence 1 from Patent FR2787121.
ACCESSION      AX025996
```


QY 1681 aggaagcagctatgaaccactccttaaaacccctgagacacatttgcattgatacat 1740
|||||
Db 1681 AGGAAGCAGCCTATGAAACCACTCTTAAACCCCTGGAAACATTGGCATTGATCATTAAT 1740
QY 1741 gctcagcacttctatgctcgaagaagcaacaacttactcaatgcgcgcg 1793
|||||
Db 1741 GCTCAGCAGCATTGTATGTGCGGAAGCAACAACTTACTCAATGCGCCGCG 1793

RESULT 3
AFPAANL 1727 bp DNA linear BCT 03-FEB-1999
LOCUS
DEFINITION A. faecalis DNA for arylacetoneitrilase, complete cds.
ACCESSION D13419
VERSION D13419.1 GI:216202
KEYWORDS arylacetoneitrilase.
SOURCE Alcaligenes faecalis (strain:JM3) DNA.
ORGANISM Alcaligenes faecalis
Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
Alcaligenes.
REFERENCE 1 (bases 1 to 1727)
AUTHORS Kobayashi,M.
TITLE Direct Submission
SUBMITTED (16-Oct-1992) Michihiko Kobayashi, Kyoto University,
Department of Agricultural Chemistry, Faculty of Agriculture;
Oiwake-cho, Kitashirakawa, Sakyo-ku, Kyoto 606, Japan
(Tel:075-753-6114, Fax:075-753-6128)
2 (bases 1 to 1727)
Kobayashi,M., Izui,H., Nagasawa,T. and Yamada,H.
Nitrilase in biosynthesis of the plant hormone indole-3-acetic acid
from indole-3-acetonitrile: cloning of the Alcaligenes gene and
site-directed mutagenesis of cysteine residues
Proc. Natl. Acad. Sci. U.S.A. 90 (1), 247-251 (1993)
93126352
COMMENT Submitted (16-Oct-1992) to DDBJ by:
Michihiko Kobayashi
Department of Agricultural Chemistry
Faculty of Agriculture
Kyoto University
Kitashirakawa-oiwakecho
Sakyo-ku
Kyoto 606
Japan
Phone: 075-753-6114
Fax: 075-753-6128.

FEATURES
source location/Qualifiers
1..1727
/organism="Alcaligenes faecalis"
/strain="JM3"
/db_xref="taxon:511"
304..1374
/codon_start=1
/transl_table=11
/product="arylacetoneitrilase"
/protein_id="BA02684.1"
/translation="MOTRKIVRAAQAQSPNYDLATGVDTKTELARQARDSCDILV
FGEITLPGYFHWLGAAPMSLKYSAKYANSISLDSAEFORIAQAARLIGIFALGY
SRRSGSLYLGGCLIDDKGOMLSRRLKPTHERYFEGEYARADLIYDTELGRVGA
LCGHEHSLPSLKYALYSQHEAHIHAMPFSLSBDAAHLSKYNMAAQIYSVEGOC
FTIASSVVTQETLDMLEVEGHNASLTKVGGSSMTFAPDGRTLAFLPHDEGLIIA
DINMEIRAPAKAINDPVGHSKPEATRLVLGHRPMTRVHSKSVIOEAEPEPHVOS
TAAPYAVSOTQSDTLVQEPS"
BASE COUNT 343 a 558 c 476 g 350 t
ORIGIN

Query Match 63.2%; Score 1132.8; DB 1; Length 1727;
Best Local Similarity 99.8%; Pred. No. 8.9e-255;
Matches 1134; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 122 tatgcagaagaataatcgctccggcagccgcgctacagcgccgctctcccaactacga 181

Db 303 TATGACAGACAAGAAAATATCTCCGGGAGCGCGCGTATACAGCGCGCTCTCCAACTACGA 362
QY 182 tctggaacaggggtgttgataaaacattgactgctgcctcgaagccgagatgaagctg 241
Db 363 TCTGGCAACGGGTGTGATATAAACCATTTAGCTGGCTCGCAAGCCCGCATAGAGGCTG 422
QY 242 tgaactgatcgtgtctgtgttaaaccttgactgcccggactacccctcgaactgcctgcctga 301
Db 423 TGACCTGATGCTGTTGGTGTAAACCTGGCTGCCCGGTATCCCTTCACAGCTCGGGCTGG 482
QY 302 cgcacccggccttgctgcgtgaataacagtlgcgcgactactatgccaactgcctcgtctga 361
Db 483 CGGACCGCGCTGCTGGCTGTAATAATACAGTGCCTCTACTATGACCAATCGCTCGCTGGA 542
QY 362 caatgcaagattcaacagcacttgcgaagccgagagacacttggatatttactgcact 421
Db 543 CAGTGCAGAGTTTCAACGCACTTGCCTCAGCGCAGCACTTGAGTATTTTCAATCCACT 602
QY 422 gggctataagcagcgagcgcgagcagccttactccttgagcgaattgcctgaatgcagaca 481
Db 603 GGGTTATAGCGAGCGCAGCGCGGCGAGCCTTTACTGCGGCATAGCTGATGACGACAA 662
QY 482 gggccagaatgctgtgtgtcgtcgtcgaataactcaaacctcaatgttgcagcgagctgt 541
Db 663 GGGCCAGATGCTGTGTGCTGCGCGTGGCAAACTCAAACTACAGTGTGAGCGCACCGTGT 722
QY 542 tggtagaaggtctagcccggaatctgatttgtctcgaacacagcctggccgctgcgtgc 601
Db 723 TGGTGAAGGTTATGCCCCGAGATGTGATTTGTCTCGACACGCACTGGCGGCTCGGTG 782
QY 602 cctgtgctgtgtagagacactgtccctctgagaagtagcgtgactccagaaga 661
Db 783 CTTGTGCTGTGGAGACACTGTCCCTTGAGCAAGTACGCGCTACTCCACGACAGA 842
QY 662 agccattcaatagccgcgcgcgcgtccttctgctgtaagaagaagcccaatgcgtc 721
Db 843 AGCATTACATTTGCGCGCTGCGCTGCTCTTTGCTGTACAGGAACAGGCCATCGGCT 902
QY 722 caagccaaagtgaacaatgctgctgcctcgaactcattcgttgaagccagctgttacc 781
Db 903 CAGCGCCAAAGTGACATGAGCTGCTGCAAACTATTGCGTTGAAGGCCACGCTTTAC 962
QY 782 catgcgcgcagcagtgctgtcaccgaagagacactgagacatgctgtgaagtgaatga 841
Db 963 CATGCGCGCCAGCAGTGTCTGTCAACCCAGAGACATGGAATCTGTGAAGTACGTAAACA 1022
QY 842 caagcctccctctgtaagaatggggcgagcagcttccatttttggccggaagcagc 901
Db 1023 CAACGCTCCCTGTAAGTGGGCGGCGGACGACTTCATGATTTTGGCCGGAAGGAGC 1082
QY 902 caattgctcccttaccctgcacagagatgacgaagccctatcatctgacgacttgaacat 961
Db 1083 CACATTGCTGCTCTACTGCGCACAGATGCGGAAGGCTGATATTTGCCGATTTGAACAT 1142
QY 962 ggaagaataatgctctgcgcgaagcgatcaagacccctgtgtgacactactcaaacccga 1021
Db 1143 GGAAGAAATTCCTTGGCCCAAGCGATCAACGACCCGTGTGGCCACTACTCCAAACCCGA 1202
QY 1022 ggcgaacccgctgtgacttgaaccttgaggacccgttgagcccaatcccggttgatccaa 1081
Db 1203 GGCACCCCTGCTGTACTGGAACCTGGGACCCCTGTGACCCATAGCTGGGTGCAATTCGA 1262
QY 1082 aagcgtgactcagaagaagcctccgagccgacgctgcaaatgactgagctgcgcctgcg 1141
Db 1263 AAGCGTATTCAGAGAAAGACTCCGACGCGACAGTGCAGAAAGTACGGCTCGCGCTGCG 1322
QY 1142 cgtcagccagactcagactcagatacgtactcgtgtgcaagaacgctcctgaccccaaaa 1201
Db 1323 CGTACGACAGACTACAGACTCGGATACGTTAGTGTTGCAAGAACGCTCGACCCCAAAA 1382
QY 1202 gatgacaagcccgagcaaatcttccggtctctgattcctctcgtgtcccgagctc 1257
|||||

Db 1383 GATGACAAGGCCCGGCAAACTGTCGGGCTTGATTCCTCTGCGTCCCGATCC 1438

RESULT 4

LOCUS E07629 1727 bp DNA linear PAT 29-SEP-1997

DEFINITION gDNA encoding nitrilase.

ACCESSION E07629

VERSION E07629.1 GI:2175764

KEYWORDS JP 1994153968-A/6.

SOURCE JP 1994153968-A/6.

ORGANISM Alcalligenes faecalis.

Alcalligenes faecalis

Bacteria; Proteobacteria; beta subdivision; Alcalligenaceae;

Alcalligenes.

1 (bases 1 to 1727)

REFERENCE

AUTHORS Endo,R., Yamada,H., Shimizu,A., Nagasawa,T. and Kobayashi,T.

TITLE PRODUCTION OF ORGANIC ACID BY TRANSFORMANT CONTAINING GENE DNA

JOURNAL CODING POLYPEPTIDE HAVING NITRILASE ACTIVITY

Patent: JP 1994153968-A 6 03-JUN-1994;

NITTO CHEM IND CO LTD

OS Alcalligenes faecalis

PN JP 1994153968-A/6

PD 03-JUN-1994

PE 17-NOV-1992 JP 1992306663

PI ENDO RYUICHI, YAMADA HIDEAKI, SHIMIZU AKIRA, NAGASAWA TORU, PI

KOBAYASHI TATSUHIKO

PC C12P7/40,C12P7/42,C12P11/00,(C12P7/40,C12P11/05),(C12P7/42,PC

C12R1/05),

PC (C12P11/00,C12R1/05);

CC strandedness: Single;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

EH key Location/Qualifiers

FT source 1..1727

FT /organism='Alcalligenes faecalis' FT

FT /strain='JM3'

FT /clone='pNMJ20'

FT 5'UTR 1..303

FT CDS 304..1374

FT /product='nitrilase'

FT 3'UTR 1375..1727

FT RBS 293..297.

FEATURES

Location/Qualifiers

1..1727

/organism="Alcalligenes faecalis"

/db_xref="taxon:511"

BASE COUNT 344 a 585 c 449 g 349 t

ORIGIN

Query Match 61.2%; Score 1097.6; DB 6; Length 1727;

Best Local Similarity 97.9%; Pred. No. 1.6e-246;

Matches 1112; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

122 tatgagaagaagaataatcgtccggcgacgcgtaagagccgctctcccaactaga 181

Db 303 TATGAGACAAGAAAATTCGTCCGGCAACCCGCTACAGCGCGCTCTCCCACTACGA 362

182 tctgcaacgagtgctgataaaccatgagctgctcgaagcccgagctgagagctg 241

Db 363 TCTGGCAACGGGTGTGATTAACAACATGAGCTGGCTCTCAAGCCCGCATGAGGGCTG 422

242 tgaactgactgctgttctggaacctgctgcgcgactatccctccacgcttgcgtg 301

Db 423 TGACCTCATCTGTTGGTCAAACTGGCTGCCGCGCTATCCCTTCACAGTCTGGCTGG 482

302 cgcacagccctgctgctggaataatagcgcgctactatgcaactgctcgtctga 361

Db 483 CGCACCGGCTGGTGGTAATAACAGTGCCTCCGCTACTATGCAACTGCTCTCGCTGGA 542

362 cagtcgagagttcaacgcatctgcacagccgcacgacacttggtatttcacgcact 421

Db 543 CAGTGACAGGTTTCAACGATTTGCCAGCGCCGACGAGCACTTGGGATTTATCATCGCACT 602

422 gggttataagcagcagcagcgagccttactcctggccaatgctgatacagacaa 481

Db 603 GGGTTATAGCGAGCGCAGCGCGGAGCCTTACTCGGSCCAATGGCTGATGACGACAA 662

482 gggcagatgctgctgctgcgcgaactcaacttaactgtttgagcgacgctgt 541

Db 663 GGGCCAGATGCTGTGGTCCGTCGTCGCAAACTCAACCTTACACATGTGTGAGCGGACGCTGTT 722

542 tggtaaggattagcccgagatcgtatgtgtgcgacacccgagcctggcgctgctg 601

Db 723 TGGTGAAGTTATGCCCGGAGATCTGATTGTGCCAGACCGAGCTGGCCGGCTGGTTC 782

602 cctgtgctgctggagagacactgtccctcttgacgaagtagcgcgtgactcccaagca 661

Db 783 CCGTGGCTCTCGAGCACCTGTCCCTTTGAGCAAGTACGGCGCTGTACTCCAGACAGA 842

662 agccattacatctgcgcctgagcgtctcttctcgtctgtaacggaacagcccatgct 721

Db 843 AGCCATTCAATTGCGCGCTGGCGCTCTTTTCGCTGTACACGCAACAGGCCCATGCGCT 902

722 cagcgccaaagtgaacatgctgctcctcgcaaatctatcgtgttgaagcgagctgttac 781

Db 903 CAGCGCCAAAGTGACATGCTGCTGCTCGCAAACTTCTATTGGTTGAAGCCACTGCTTTC 962

782 catcgccgcgaagcagctgtctgtaacccaagagacactgacatgctgtgaagtgtgaaca 841

Db 963 CATCGCGCGCAGACAGTGTGTACACCAGAGACACTGACATGTGTGAAGTGTGAACA 1022

842 caagcctccctgctggaagtggtggcgcgagcttccatgatttttggcgcggaagagc 901

Db 1023 CAACGCTTCCTGCTGAAAGTGGGGCGGCGGAGTTCCATGATTTTGGCCCGGAGGAGG 1082

902 cacattgctccttaccctgcccacacagatgcgaaagcctgcatcttgcgactgaat 961

Db 1083 CACATTGGCTCCTCTACTCTGACACAGATGCCAAGGCCCTGATCATTTGCGGATCTGAACAT 1142

962 ggaagaaltgctcttcgcgaagcgatcaacgacccctgtggccaactactcaaaccca 1021

Db 1143 GGAAGAAATTTGCTTCGCAAGGCGATCAACGACCTGTGGGCCACTACTCCAAACCCGA 1202

1022 ggcacccgctcgtgactgagacttgggacccgttagccatgactcgggtgacttcca 1081

Db 1203 GGCCACCGCTGTGACTGAGCACTGGGGCACCGTAGCCCATGACTCGGGTCAATTCNA 1262

1082 aagcgtgataccagaagaagctcccgagcgcagctgcaaaatagcgtgcgcctgcgc 1141

Db 1263 AAGCGTGAATCCAGCAAGAACCTCCGAGCGGCACTGCAAACTAGCCCTGCCCGCTGCC 1322

1142 cgtcagcagactcagagactcgatacgtactggtgcaagaacgctctgaccccaaaa 1201

Db 1323 CCTACGCCACACTCAAGATGCGATACGCTACTGGTGAAGAAGCGTCTGTGACCCCAANA 1382

1202 gctgacaagcccggaacactgctcgggtcttgattctcttctgcttcccgatcc 1257

Db 1383 GATGACAAGGCCCGGCAAACTGTCGCCCTTCACTCTTGTGCTCCCGATCC 1438

RESULT 5

LOCUS E07629 7508 bp DNA linear BCT 15-APR-1994

DEFINITION E.coli RNA operon (rnb) coding for Glu-tRNA-2, 5S, 16S and 23S rRNA.

ACCESSION J01695

VERSION J01695.1 GI:147581

KEYWORDS

SOURCE Escherichia coli DNA.

ORGANISM Escherichia coli

Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1518 to 3059)

```

AUTHORS      Brosius,J., Palmer,M.L., Kennedy,P.J. and Noller,H.F.
TITLE        Complete nucleotide sequence of a 16S ribosomal RNA gene from
JOURNAL      Escherichia coli
REFERENCE    Proc. Natl. Acad. Sci. U.S.A. 75, 4801-4805 (1978)
AUTHORS      2 (bases 811 to 1543)
JOURNAL      Georgas-Toth,E., Boros,I. and Venetianer,P.
REFERENCE    79116209
TITLE        Structure of the promoter region for the rnb gene in Escherichia
JOURNAL      coli
REFERENCE    Nucleic Acids Res. 7, 2189-2197 (1979)
AUTHORS      3 (bases 3500 to 6403)
JOURNAL      Brosius,J., Dull,T.J. and Noller,H.F.
REFERENCE    80101075
TITLE        Complete nucleotide sequence of a 23S ribosomal RNA gene from
JOURNAL      Escherichia coli
REFERENCE    Proc. Natl. Acad. Sci. U.S.A. 77, 201-204 (1980)
AUTHORS      4 (bases 1 to 7508)
JOURNAL      Brosius,J., Dull,T.J., Sleeter,D.D. and Noller,H.F.
REFERENCE    80145594
TITLE        Gene organization and primary structure of a ribosomal RNA operon
JOURNAL      from Escherichia coli
REFERENCE    J. Mol. Biol. 148, 107-127 (1981)
AUTHORS      5 (bases 6412 to 6507; 6603 to 6657)
JOURNAL      Singh,B. and Apirion,D.
REFERENCE    82055570
TITLE        Primary and secondary structure in a precursor of 5S RNA
JOURNAL      Biochim. Biophys. Acta 698, 252-259 (1982)
REFERENCE    6 (bases 6518 to 6931)
AUTHORS      Sarmentenos,P., Sylvestre,J.E., Contente,S. and Cashel,M.
REFERENCE    differential stringent control of the tandem e. coli ribosomal rna
AUTHORS      promoters from the rna operon expressed in vivo in multicopy
TITLE        plasmids
JOURNAL      Cell 32, 1337-1346 (1983)
REFERENCE    83180429
AUTHORS      7 (bases 3380 to 3509; 6392 to 6503)
JOURNAL      King,T.C., Slideshmuah,R. and Schlessinger,D.
REFERENCE    RNase III cleavage is obligate for maturation but not for function
AUTHORS      of Escherichia coli pre-23S rRNA
TITLE        Proc. Natl. Acad. Sci. U.S.A. 81, 185-188 (1984)
JOURNAL      84119433
REFERENCE    8 (sites)
AUTHORS      Gross,A., Boros,I. and Venetianer,P.
TITLE        Analysis of the complex transcription termination region of the
JOURNAL      Escherichia coli rnb gene
REFERENCE    Eur. J. Biochem. 201, 653-659 (1991)
AUTHORS      9 (sites)
JOURNAL      Prestle,J., Schoenfelder,M., Adam,G. and Mundry,K.W.
REFERENCE    type 1 ribosome-inactivating proteins deurate plant 25S rRNA
AUTHORS      without species specificity
TITLE        Nucleic Acids Res. 20, 3179-3182 (1992)
JOURNAL      92319651
FEATURES
source       location/Qualifiers
            1. 7508
            /organism="Escherichia coli"
            /db_xref="taxon:562"
            1. 250
            /note="putative VECTOR sequence lambda; putative"
            misc_difference 1068..1072
            /citation=[2]
            misc_signal 1212..1218
            /note="promoter: RNA polymerase binding site"
            misc_rna 1223..>3306
            /note="rnb precursor RNA (alt.)"
            misc_signal 1331..1337
            /note="alternate promoter"
            misc_rna 1344..>3306
            /note="rnb precursor RNA (alt.)"
            misc_difference 1507
            /citation=[2]
            /replaceme=""
            rRNA 1518..>3059

```

Query Match	Best Local Similarity	Score	DB	Length
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0;	97.1%;	Pred. No. 1.6e-106;		
trNA	/note="16S rRNA (rnbB)"	3231..3306		
trNA	/note="codon recognized: GAA: Glu-trNA-2 (rnbB)"			
trNA	/product="rRNA-Glu"	3500..6403		
trNA	/note="23S rRNA (rnbB)"	6496..6615		
BASE COUNT	1699 a 1694 c 2166 g 1679 t			
ORIGIN	1 bp upstream from BamHI site.			
Query Match	27.9%;	Score 500;	DB 1:	Length 7508;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0;	97.1%;	Pred. No. 1.6e-106;		
Db 1267	agagtcgacccctcaagcatcgaagcttggtccaccctgaaccatgcccgaatcagaag	1326		
Db 6430	AGAATTTCGCTTGCGGCGCAGTACCGGGGTGTCTCCACTACCCCATGCCGAATCAGAA	6549		
Db 1327	tgaaagccgtatgcgcgaatgtagtgggtgtcctcccatgagagtaggaaatgccc	1386		
Db 6550	TGAACCGCGGTGAGCGCCCATGCTAGTGTGGGTCTCCCATCGAGAGTAGGAAGTCC	6609		
Db 1387	agagctcaataaagaaagagctcaagtcgaagaagctggccttcgtttatcgttgt	1446		
Db 6610	AGGCATCAAAATTAAGAAAGAAAGCTCAGTCGAAGACTGGCCCTTTCGTTTATCTGTGT	6669		
Db 1447	ttgtcgtgaagcgtctcctcgtatgtaggaacaaatccgcgcgagagcgaattgaacgttcg	1506		
Db 6670	TTCGTGGTGAACGCTCTCTCTGTAGTGAACAATTCGCGCGGAGCGGATTTGAACGTTCG	6729		
Db 1507	aagcaacgccccggaaggtgctgcgcgaagagccgcgaataaactgcaagcatcaaat	1566		
Db 6730	AAGCAACGCGCGCGGAGGGGTGGCGGAGACGCCCATTAACATCCAGCATCAAAAT	6789		
Db 1567	aagcgaagagccatcctgaagctgtgcttttgtgtttcacaacatcttcctgtgtc	1626		
Db 6790	AAGCGAAGGCGCATCTCAGCGGATGGCTTTTGGCTTCAACAACTTCTGTCTGT	6849		
Db 1627	atatcacaagcatcccccacagaatacgtgaactagcctgttttgatcagaagaa	1686		
Db 6850	ATATCTACAGCATCCCTCCCAACATACGTAACCTACCTCGTTTGGCATCAGAA	6909		
Db 1687	gcagcatgaaacacatccttaaaacccctggaacacatcttgcatgatcataatgctcag	1746		
Db 6910	GCAGCTATGAACCATCTCTTAACACCTGGAACACATTTGGCATGATCATATCTCTAG	6969		
Db 1747	cacatctgatgtgcccgaagaagcaacaacatctcaatgcccgc	1790		
Db 6970	CACATTTGATGTGGCGGAAGACGACAAACATTAATCATGCTCTG	7013		
RESULT 6				
ECRRNBZ	7508 bp	linear	BCT 06-MAY-1997	
LOCUS				
DEFINITION	E. coli ribosomal operon rnb encoding the 16S ribosomal RNA. Also transfer RNA specific for Glu, 23S ribosomal RNA and two unidentified open reading frames. This sequence was obtained from the transducing phage lambda-rlif-d 18 (BAMHI fragment).			
ACCESSION	V00348			
VERSION	V00348.1			
KEYWORDS	16S ribosomal RNA; 23S ribosomal RNA; ribosomal RNA; transfer RNA; Glu; unidentified reading frame.			
SOURCE	Escherichia coli.			
ORGANISM	Escherichia coli			
REFERENCE	Bacterla; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.			
AUTHORS	1 (bases 1 to 7508)			
TITLE	Broslus J., Dull T.J., Sleeter D.D. and Noller H.F.			
JOURNAL	Gene organization and primary structure of a ribosomal RNA operon from Escherichia coli			
	J. Mol. Biol. 148 (2), 107-127 (1981)			

JOURNAL REFERENCE AUTHORS TITLE	82055570 2 (bases 1 to 7508) van Keulen,H., Gutell,R.R., Gates,M.A., Campbell,S.R., Erlandsen,S.L., Jarrold,E.L., Kulda,J. and Meyer,E.A. Unique phylogenetic position of Diplomonadida based on the complete small subunit ribosomal RNA sequence of Giardia ardese, G. muris, G. duodenalis and Hexamita sp
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	PASEB J. 7 (1), 223-231 (1993) 93138311 3 (bases 1 to 7508) Weltmann,C.J., Cunningham,P.R., Nurse,K. and Ofengand,J. Chemical evidence for domain assembly of the Escherichia coli 30S ribosome
JOURNAL MEDLINE AUTHORS TITLE	PASEB J. 7 (1), 177-180 (1993) 93138304 4 (bases 1 to 7508) Ringquist,S., Cunningham,P., Weltmann,C., Foremanoy,L., Pleij,C., Ofengand,J. and Gold,L. Translation initiation complex formation with 30 S ribosomal particles mutated at conserved positions in the 3'-minor domain of 16 S RNA
JOURNAL MEDLINE AUTHORS TITLE	J. Mol. Biol. 234 (1), 14-27 (1993) 94047056 5 (bases 1 to 7508) Sako,Y., Takai,K., Ishida,Y., Uchida,A. and Katayama,Y. Rhodococcus obakensis sp. nov., a modern lineage of extremely thermophilic marine bacteria
JOURNAL MEDLINE COMMENT FEATURES	Int. J. Syst. Bacteriol. 46 (4), 1099-1104 (1996) 97016816 On May 8, 1997 this sequence version replaced gi:42877. Location/Qualifiers 1..7508 /organism="Escherichia coli" /db_xref="taxon:562" 240..254 /note="lambda attachment site" 275..1144 /note="unidentified reading frame 1" /codon_start=1 /transl_table=1 /protein_id="CAA23638.1" /db_xref="GI:42878" /db_xref="SWISS-PROT:P22634" /translation="MKRSMAKKLDDGNPCLAATPSEDRPVLVDSGVGLSYDEI RHLPLDLHYIAFDNVAEPYGEKSEAFIVERVNAIVAVQRYPLALAVACNASTV SLPLRKPFDPYGVVPAIKPARLRTANGVGLATRGTVKRSYTHEILARFNECC IEMQSAENVEVLAAPKLHGEDVSDALAKRIIRPMVRKMEPDVTVLGCTHPLLOE LOVPEGRVLDSGAALAIRRTAWLLEHRAEPKASDANIACMAMTPCAEQLLVLR YGPETLEKLYAVG"
rRNA	1518..2959 /product="16S ribosomal RNA"
tRNA	3231..3306 /note="transfer RNA - Glu - 2"
rRNA	3500..6403 /product="23S ribosomal RNA"
BASE COUNT	1970 a 1694 c 2166 g 1678 t
ORIGIN	
Query Match	27.9% Score 500; DB 1; Length 7508;
Best Local Similarity	97.1% Pred. No. 1,6e-106;
Matches 509; Conservative	0; Mismatches 15; Indels 0; Gaps 0;
1267	agagtcgacgcctgcgaagcctgcgaagcttggtgtccaccctgcgaccacatgcgcgaactcagag 1326
6490	AGAAATTTCTCCCTGCGCGCGACGTAACCGGCGTGTCCACCTGACCCCAATGCCGAAGCTGAGAAG 6549
1337	tgaaagccgctgagcgcgcgaatgtagtctgtggtgtctcccatctgcgagagtagagaaactgcc 1386
6550	TGAAAGCGCGTAAGCCGCCGATGTAATGTGGGTCTCCCATATCGAAGATGAGGAAGTAAGTGGC 6609
1387	agggcatcaataaagaagaagctcagtcgcgaagaagctggcgcttcgtttatctgttct 1446

Db	Accession	Source	Organism	Reference	Authors	Title	Journal	Year	Volume	Page	DOI	PMID
Db	6610	AGGCATCAATPAALAAAGCGAAGGCTCAGTCGAAAGACTGGCCCTTCGTTTATCTGTGT	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					6669
Oy	1447	ttgtcgtgtaacgcctctcctgaatgagacaatcccgcgagcgatttgaacgttcg	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					1506
Db	6670	TTGTGCGTAACGCTCTCTCTAGTAGGACAAATCCGCCGGGAGCGGATTTGACGCTTCCG	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					6729
Oy	1507	aagcaacggcccggaagggtggcgggcagaacgcccgcataaactgcgaagcatcaatt	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					1566
Db	6730	AAGCAACGCCCGGAGGGGTGGCGGGCAGACGCCCGCATTAACATGCCAGCATCAATAT	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					6789
Oy	1567	aagcgaagggcatctctgaagatgagccttttggctttcacaacctctccgtgtc	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					1626
Db	6790	AAGCAAGAGCCATCTCAGCGATGGCCCTTTTGGCTTTCACAAACCTTCCTGTCTC	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					6849
Oy	1627	atatataaagccatccccccaagatagcgtlaaactagcctgcttttgatcagagaa	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					1686
Db	6850	ATATCTACAAGCCATCCGCCACACATACGGAATACCTACCTCTTTTGCATCAGGAA	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					6909
Oy	1687	gcacgtatgaacacactccttaaaacccctggaacacatttggcatgcatcaatgctcag	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					1746
Db	6910	GCAGCTATGAACCACTCTTAAACCTCGAACAACATTTGGCATGATCATATATGCTCAG	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					6969
Oy	1747	caacattgatatgcccgaagacgaagaacaaattaccatgccc	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					1790
Db	6970	CACATTTGATGTGCGGAGAGCAACACAAATTTACTCAATGACCTCG	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					7013
RESULT 7												
AE000471												
LOCUS	AE000471	000096	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
DEFINITION	AE000471	000096	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
ACCESSION	AE000471	000096	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
VERSION	AE000471.1	GI:1790404	Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
KEYWORDS			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
SOURCE			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
ORGANISM			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
REFERENCE			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
AUTHORS			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
TITLE			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
JOURNAL			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
REFERENCE			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
AUTHORS			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
TITLE			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
JOURNAL			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
REFERENCE			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
AUTHORS			Escherichia coli K12	11679 bp	DNA	linear	BCT 01-DEC-2000					
TITLE			Escherichia coli K12	11679 bp	DNA	linear</						

frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 [e-mail: mark@ember.gatech.edu]. Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (<http://cgsc.biology.yale.edu>). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (<http://www.genetics.wisc.edu>). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

FEATURES

source

1..11679

Location/Qualifiers

/organism="Escherichia coli K12"

/strain="K12"

/sub_strain="MG1655"

/db_xref="taxon:83333"

224..2068

/gene="btub"

/note="b3966"

224..2068

/gene="btub"

/function="membrane: Outer membrane constituents"

/note="o614: 100 pct identical amino acid sequence and equal length to BTUB_ECOLI SW: P06129; CG Site No. 946"

/codon_start=1

/transl_table=11

/product="outer membrane receptor for transport of vitamin B12, E colicins, and bacteriophage BF23"

/protein_id="AAC76948.1"

/db_xref="GI:1790405"

/translation="MIKKSILITACSVTFSAADOTSPDILVYFANREPRSTVLA

PTVYVRODIDRMOSTVNDVLRRLPGVDITONGSGQLSIFRTGNSHVLVLDG

VRNLNAGVSGADLSQPIALVQREYIRGPSAVYSDAIGVNIITTRREPGEI

SAGWSNSYQNDVSTOQQLGDKRTVLLGDYAHGVDVAVGNATGTAQDNDFL

STLYGALHNFTDAMSGVRGVDYNDRTNDAVYSPGSLDTRKLYSOSMDAGRY

NGLIKSOLITSYSHSKDYNVPHYGRVDSATIDEMKOYTWOMANNVYVGHSGTSG

VMOKOTPRGTVGVEDYGDONTGILYGLCOOQGFPEGAARSDNQGRCRGTWQ

TSAGMEFEGYRFIASYGTSTYKAPNLGQLIGFGNPNLDEPKSOMEAFEGYLVAGN

WRISGRVNDVLDIDVDTLRYNMGKARIGVFATNFDGPIHTVYSYDVAGN

ATDTPLLRRARQOVKYOLDMQLYDFDMGITVQYLGTRDKDYSSYPYTVMGVSL

MDLAVALPYTSHLTVAGKIANLDFDKDYEYVGYQFAGREYTLGSYTF"

2001..2870

/gene="muri"

/note="b3967"

2001..2870

/gene="muri"

/EC_number="5.1.1.3"

/function="enzyme: Murein sacculus, peptidoglycan"

/note="o289: 100 pct identical to 285 amino acids of MUR1_ECOLI SW: P22634 but contains 4 additional N-term

residues; alternate gene name dga"

/codon_start=1

/transl_table=11

/product="glutamate racemase, required for biosynthesis of

D-glutamate and peptidoglycan"

/protein_id="AAC76949.1"

/db_xref="GI:1790406"

/translation="KROSMATKLDQGNPCLATSPSEPPFVLVDPDSGVGSLSVYDEI

RLPLDPIHYAFDVAAPYGESEAFIVERVAIVTAQOERPLALAVACNASTV

SLPALREKDFPDVGVPAIKPAARLTANGIVGLATRGTVKRSYTHELIARPANCQ

IEMLSAEVLEAEKIHGEDVSLDAKRLIRPLRMKPEPDTVLGCTHFPLOBEL

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

rRNA

LOVLEPSTRLVDSGAIAIRRTAWLLEHAPDAKSDANIAFCAMTPGAEOLLPVLQR

YGEPTLEKLAIVLC"

3244..4785

/gene="rrsB"

/note="CG Site No. 188; 16S rRNA of rrnB operon"

/product="16S ribosomal RNA"

/function="RNA: Ribosomal and stable RNAs"

3244..4785

/gene="rrsB"

/note="b3968"

4957..5032

/gene="gltT"

/note="anticodon: UUC; CG Site No. 680; in rrnB operon;

alternate gene name tglB"

/product="tRNA-Val"

/function="RNA: tRNA"

4957..5032

/gene="rrlB"

/note="CG Site No. 202; 23S rRNA of rrnB operon"

/product="23S ribosomal RNA"

/function="RNA: Ribosomal and stable RNAs"

5226..8129

/gene="rrlB"

/note="b3970"

8222..8341

/gene="rrfB"

/note="CG Site No. 209; 5S rRNA of rrnB operon"

/product="5S ribosomal RNA"

/function="RNA: Ribosomal and stable RNAs"

8222..8341

/gene="rrfB"

/note="b3971"

8642..9670

/gene="murB"

/note="b3972"

8642..9670

/gene="murB"

/EC_number="1.1.1.158"

/function="enzyme: Murein sacculus, peptidoglycan"

/note="o342: 100 pct identical to MURB_ECOLI SW: P08373"

/codon_start=1

/transl_table=11

/product="UDP-N-acetylenoipyruvoylglycosamine reductase"

/protein_id="AAC76950.1"

/db_xref="GI:1790407"

/translation="MNHSLKPMWTFEDIDNMAOHVCAEDPQOLNANQVATAGRPVL

ITGEGSNVLEEDYRGTVLINRKIGLEIHDEPAMVLIHAGAGNHRILVYTLQSGMP

GLENLALIGCVSSPIONIGAVGELGVCAVVDSEVLATGKVRLTAKERFGYRD

SIFKEHYORFAIVAVGLRPLKPEMOPVLVGLDTRDPTTVPOQVFNAVCHMRTKL

PPKRYNAGSFEKPNVSAETAKALLSOPFAPARNPQADGSYKLAAGWILQOCOLKG

MOIGSAVHRQOALVLINEDNAKSEDDVGLAHNVHQKVEKFWLMEPRVFRIGASGE

VSAVERIIS"

9667..10632

/gene="bira"

/note="b3973"

9667..10632

/gene="bira"

/EC_number="6.3.4.15"

/function="enzyme: Biosynthesis of cofactors, carriers:

Biotin"

/note="o321: 100 pct identical to BIRA_ECOLI SW: P06709;

CG Site No. 952; alternate gene name bior"

/codon_start=1

/transl_table=11

/product="biotin-[acetylCoA carboxylase] holoenzyme

synthetase and biotin operon repressor"

/protein_id="AAC76951.1"

/db_xref="GI:1790408"

/translation="MKDNTVPLKIALALLANGFPHSGEOLGFTLGSRAAIINKHIOTLR

D

K

G

S

V

L

P

E

D

T

R

I

Q

L

I

N

K

O

T

L

I

D

E

R

S

V

L

P

E

D

T

R

I

Q

L

I

N

K

O

T

L

I

D

E

R

S

V

L

P

E

D

T

R

I

Q

L

I

N

K

O

T

L

I

D

E

R

S

V

L

gene	/gene="coa"		/note="b3974"		complement(10661..11611)	
CDS	/gene="coa"		/EC-number="2.7.1.33"		/function="enzyme; Biosynthesis of cofactors, carriers:	
Query Match	27.9%	Score 500:	DB 1:	Length 11679:		
Best Local Similarity	97.1%	Pred. No. 1,7e-106:				
Matches 509:	Conservative 0:	Mismatches 15:	Indels 0:	Gaps 0:		
QY 1267	agagtcgacccgcgaaggatcgtcaagctcttgggtcccaaccctgaccacatgctgcagactcagaag	1326	DB 8216	AGAAATTTCCTCGGGCGCGAGTACGGCGGTGTCCACCTGACCCCAATGCGCAATCAGAG	8275	
QY 1327	tgaacgcccgtcagccgcatgttagtgttggtgtctccccaatgagaaatagaactgcc	1386	DB 8276	TGAACGCCGTAAGCGCCCATGTGTAGTGTGGGGTCTCCCATAGCAGAGATAGGAACATGCC	8335	
QY 1387	aggcatcaataaaacgaaaggtctcagtaagaactgtggccttctgttatactgttgt	1446	DB 8336	AGGCATCAAAATAAAAAGAAAGCTCAGTCGAAGAAGCTGGCCCTTTCGTTTATCTGTGT	8395	
QY 1447	ttgtcgtggaacgcctctctctagtagagaacaatccgcggagagcgagattgaagttgcg	1506	DB 8396	TTGTGCTGGAACGCTCTCTCTAAGTAGAGCAAAATCCGCCGGAGCGGATTTGAACCTTCCG	8455	
QY 1507	aagcaacgcccgcgaggttgccggtgcgaagacgcccgcataaaactgtccagcgatcaaat	1566	DB 8456	AAGCAACGCCGCCGAGGGTGGCGGGCAGAGACGCCGCATTAACGTCCAGCGCATCAATT	8515	
QY 1567	aagcaagaagcgcctctcgtcagcgatgtgccttlttcglttcaaaaactctctcgtgc	1626	DB 8516	AAGCAAGAGCCCATCTCAGCGGATGGCTTTTGGCGTTTACAACAACCTTCTCTGCTC	8575	
QY 1627	atactacaagccatcccccaagatacgtttaactcgttcgttlttgatacggaa	1686	DB 8576	ATATTTACAAAGCCATCCCCCACACATACGGTAACCTAGCCGTGTTTGCACTCAGGAA	8635	
QY 1687	gcagctcgaacacatctcttaaaacccctggagacacatttgcattgatcataatgtcag	1746	DB 8636	GCAGCTATGAMCCATCTCTTAAACCCCTGGAGAACACATTTGGCATTTGATATATCTCTCAG	8695	
QY 1747	cacattgtatgtgcccgaagaacgaacaataactactcaatgccc	1790	DB 8696	CACATTGTATGTGCCGAAGACGAAACACAAATTAATCAATGCTG	8739	
RESULT 8						
CVU72488/c	CVU72488	11918 bp	DNA	circular	SYN 23-NOV-1996	
LOCUS						
DEFINITION	Cloning vector pRNAB, complete sequence.					
ACCESSION	U72488					
VERSION	U72488.1					
KEYWORDS	GI:1684862					
SOURCE						
ORGANISM	Cloning vector pRNAB.					
REFERENCE	Cloning vector pRNAB					
AUTHORS	artificial sequence; vectors.					
TITLE	1 (bases 1 to 11918)					
JOURNAL	Lee,K., Holland-Staley,C. and Cunningham,P.R.					
REFERENCE	Genetic analysis of the Shine-Dalgarno interaction: selection of					
AUTHORS	alternative functional mRNA-rRNA combinations					
TITLE	Unpublished (1996)					
JOURNAL	2 (bases 1 to 11918)					
REFERENCE	Lee,K., Holland-Staley,C. and Cunningham,P.R.					
AUTHORS	Direct Submission					
TITLE	Submitted (24-SEP-1996) Biological Sciences, Wayne State					
JOURNAL	University, 5047 Guillen Mall, Detroit, MI 48202, Wayne					

FEATURES		Location/Qualifiers
Source	1..11918	/organism="Cloning vector pRNA8" /specific_host="Escherichia coli" /db_xref="taxon:54395" complement(<190..876) 190..849 /gene="cam" complement(190..849) /function="confers chloramphenicol resistance" /codon_start=1 /transl_table=11 /product="chloramphenicol acetyl transferase" /protein_id="AAB36547.1" /db_xref="GI:1684863" /translation="MEKKITGYTVDISQWHRKEHFAFOSVAQCTYNTQYVOLDTAFELKTYKKKKHKFYPAFHILHARLNMAHPEFRAMKMDGELVIMDSVPCYTFEHOETEFSLMSEYHDDPROFLHITYSDQVACYGENLAIFRFGIENMFPSANPWSTSDLVVAAAMNFAFAPFTMKRYTYGDKVLMPALIOVHNAVCDFHGRMLNELOOTIDEMOGGAA"
CDS		
RBS		
CDS		
rRNA		
rRNA		
tRNA		
rRNA		
misc-feature		
CDS		
BASE COUNT	2730 a 3372 c 2794 g 3022 t	
ORIGIN		
Query Match	27.9%: Score 500; DB 12; Length 11918;	
Best Local Similarity	97.1%: Pred. No. 1.7e-106;	
Matches 509; Conservative	0; Mismatches 15; Indels 0; Gaps 0;	
Qy	1267	aaagtcgacctgcagcgatcgaaaccttgggtccaccatgcacatgcacgaactagaag 1326
Db	5499	agaatttgcctggcgacatgacgcggctgtgctccacctgaccccatgcccgaactcagaag 5440
Qy	1327	ttaaacgcctgtagcgccagatgtgaatgtaggtctccccaatcgagagatagaaggaactgcc 1386
Db	5439	tgaaacgcctgtagcgccagatgtgaatgtaggtctccccaatcgagagatagaaggaactgcc 5380

QY	1387	aggc tcaaataaacgaaggctcag tcaagaagcttggccttcggttttcacgtgt	1446
Db	5379	AGGCATCAAAATTAAGAACGAAGGCTCAGTCGAAGACTGGGCGCTTTCGTTTATCTGTTCT	5320
QY	1447	ttgtcgtgtgaacgctctccttgagtaggaacaatccgcgcggagcgagatttgaaacttgcg	1506
Db	5319	TTGTCGGTGAACGCTCTCTCTGATGAGGAACAAATTCGGCCGGAGCGGATTTGAAACGTTGGG	5260
QY	1507	aagcaacggcccggaaggttgcgcggcagaagccgcgcaatactgcagcagcaatc	1566
Db	5259	AAGCAACGGCCCGGAGGGGCGGGCAGGACCCCGCCCTAAACATCGCCAGGCATCAAAAT	5200
QY	1567	aagcgaagagccaatcctgaacggaatggccttttgcgtttctacaacatcttcgtgcgc	1626
Db	5199	AAGCAAGAGGCATCTCTGACGGATGGCGCTTTTGTGGTTTCTACAAACTTCTCTGTGCTC	5140
QY	1627	atatataagcatalcccccacacagataacgtaaacatagcctcgtttttgcatacgaana	1686
Db	5139	ATATCTACAGGCATCCGCCACAGATAGGTAACACTAGCTCGTTTGTGCATCAGGANA	5080
QY	1687	gcagctatgaacactcctctaaacccctggaacacatlttgcatgataataatgctcag	1746
Db	5079	GCAGCTATGAACCACTCTTAAACCTCGGAACACATTTGGCATGTGATCATATGCTCAG	5020
QY	1747	caactctgatctgtccgaagacgaacaaacaaatattacatgaagccgc	1790
Db	5019	CACATTGTATGTGCGGAAGACGACACACATTAATCAATGAAGCCTG	4976

RESULT	9
ECOW89	176195 bp DNA linear BCT 17-DEC-1993
LOCUS	ECOW89
DEFINITION	E. coli chromosomal region from 89.2 to 92.8 minutes.
ACCESSION	U00006
VERSION	U00006.1 GI:409785
KEYWORDS	Escherichia coli (sub-strain MG1655, strain K-12) (library: lambda)
SOURCE	

ORGANISM	Escherichia coli Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
REFERENCE	1 (bases 1 to 176195)
AUTHORS	Blattner, F.R., Burland, V., Plunkett, G. III, Sofia, H.J. and Daniels, D.L.
TITLE	Analysis of the Escherichia coli genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes
JOURNAL	Nucleic Acids Res. 21 (23), 5408-5417 (1993)
MEDLINE	94089392
REFERENCE	2 (bases 1 to 176195)
AUTHORS	Blattner, F.R.
TITLE	Direct Submission
JOURNAL	Submitted (03-SEP-1993) 608-263-7459
COMMENT	On Oct 29, 1993 this sequence version replaced g1:396288.

FEATURES	Location/Qualifiers
SOURCE	1. .176195

```

1. 1/6195
   /organism="Escherichia coli"
   /strain="K-12"
   /sub_strain="MG1655"
   /db_xref="taxon:562"
   /notes="This sequence comprises the following lambda
clones: EC18-126, EC14-54, EC21-52, EC17-24, EC17-203,
EC19-105, EC18-208, EC18-110, EC17-200, EC18-33, EC30-262,
EC30-41, EC30K637-2a, EC30K637-5, EC27-975, EC22-27;
lambda or Janus vectors were used for subcloning"

```

```
gene      1. .1257      /gene="katg"
misc_feature <1. .1737      /note="corresponds to M21516; ECOKATGA(1225. .2805)"
CDS       <1. .1257
```

```

/standard_name="catalase HPI; catalase-peroxide; catalase
/EC_number="1.11.1.6"
/note="CG Site No. 14983"
/codon_start=1
/transl_table=11
/product="catalase hydroperoxidase I"
/protein_id="AAC43048.1"
/db_xref="GI:396289"
/translation="AITSGLVWVQTPTQWMSNYFEENLTFKEMWQTSPPAGAIQFEF
VDAPEIPEPDPSPKKRRPTMLYDILTRDFEFKISRRLNDQAAENEFARKWRF
LTHRMGSPRSRTIGPEVREDDIWDPLPTPTNTEDDIILDKPAINDSLSELY
SVAAASAFIRGGRDGRGANGARLALMPQRPDVAANVRAFLPVLEKIQKSGKASL
DIIVLAGVVGVEKASAAGLSIHPVPAAGRYVARDDQDIEMFELLEFIAGFRYR
RDLSTVEELKIDRAKOQLTTPAEMTALVGGVARGANFDSKNGVFPDRVGLSND
FNLLIDMRKMAKTDESKELPEGRGRFETGEVFTASRDALVFGSSVLAVAEYVAS
DAHEKFPVDAVAAMKVNMLDRFDLL"
misc_feature
1..93

```

```
misc_feature
    /note="93 bp overlap with end of l19201 (ECOMU87)"
    /pos=161
    /gene="katG"
misc_feature
    /note="corresponds to lambda clone EC18-126"
    /pos=156..15760
misc_difference
    /note="corresponds to lambda clone EC14-54"
    /pos=938..11478
```

	/note="G in M21516; c here"	.
terminator	1266.	.1305
	/note="putative"	
CDS	1317.	.2255

```

//note="similar to Desulfurolobus ambivalens hypoth. 28.3
kDa protein in sor 3' region"
//codon_start=1
//transl_table=11
//label=ORF_0312
//protein_id="AAC43049.1"
//db_xref="GI:409786"
//translation="MGERKLIHTGSLMSAAGKSNPLAIGLVLTPLTWSYIWMFKQVY
SYIGAFDALRCIRGALVLFVILLRGRMPPEPKYTLIALLOTGCGVGLAKOMALAI
VSGGAKVALIESTMPFWVYFAALFGEIRRGQYFALIAEGLFLVLPOMLDEP
SKKSMALITISVSGCAIYAKRLXYAHPRIIDLSTSWOMLYAAALMVSVALIPDI
REIDMPTPERKLAYSILATIALALMSLWLFYIKNLPASIASLSTLAVPGCVGLTSMRI
LGENGAAVGSGIVLILVALAVSRKKEAVSVKRI"
1429..1432
hsc_difference

```

```

/note="C in M21516; G here"
complement(2882..2899)
/codon_start=1
/label_table=1
/label_id="AAC43050.1"
/protein_id="AAC43050.1"
/db_xref="GI:396291"
/translation="MKASIALLSITAFPSHSLSKSPAVPTVVOALQANTLADMGGA
OITSEITPYDAVYVOLTYPGCGVPOEBCVSPVYIATLRASQVLDOKLYVHDMADK
EIPQWKLKRPDSNIDHRRVPLETWEFSRHDKTPTSKNSPDYQAGDVISRLDGLA
HIGVYSDGFARDGTPLVLIHNIGAGAQDEEDVLFNMRMVGHYRFPYK"
complement(2886..2915)

```

```
misc_structure      2912.3053  
/note="predicted bend of 81 degrees"  
repeat_region      3031.3064  
/standard_name="REP; repetitive extragenic palindromic  
element"  
misc_structure      /note="contains 1 REP sequence"  
3090..3219
```

```
gene      /note="predicted bend of 75 degrees"
           complement(1174..4316)
           /gene="g1da"
           complement(1174..4316)
CDS       /gene="g1da"
           /note="similar to Bacillus stearothermophilus glycerol
           dehydrogenase"
           /codon_start=1
           /transl_table=11
           /label_ORF_f380
           /product="glycerol dehydrogenase"
           /protein_id="AAC43051.1"
           /db_xref="GI:396292"
           /translation="MPHLALLSKGAIMDRIIOSPKYIOGADVIRLGEYLKPLXER
           WLVDGDKPVLGFAOSTVEKSEFKDAGIVYIAPGEGSCONEIDRLGIAETAGCAIL
           GIGGSKTIDTAKALAHFMKVPYAIPTASTAPSCALSIVITDEEPRRIILLRNP
           NVNIVDTKIVAGAPARLAAIGDALATWFEARACSRGATTMAGCKTOALALAEI
           CYNLTLEEGEKAMLAEOHVTPALEREIVTAANTYLSGVESGGLAAHAHNGILTAI
           PDAAHYHGEKVAFGTLTQVLVLENAVEIEITVAALSHAVGCLPTLAOLDIKEDVPAK
           MRIVAAACACEGETIHNMPGATPDQVYVALLVADQYGGRFIOEME"
misc_structure
           4022..4156
           /note="predicted bend of 75 degrees"
           complement(4288..4950)
CDS       /note="similar to Bacillus subtilis hypoth. 20 kDa
           protein, in tsr 3' region"
           /codon_start=1
           /transl_table=11
           /label_ORF_f220
           /protein_id="AAC43052.1"
           /db_xref="GI:396293"
           /translation="MELYLDITANVAEVERLARIPIAGVTNPSSIJAASKESIWEVLP
           RLQKAIIGDEGILFAOTMSRDAGQWEEARLDAIPGIYVKIIPVSEGAALKIKKE
           GTTGTATVYSAOGILLALAGAKYVAPVNVVDAGGSGRTVOQLLLEMHAPES
           MYLAASFTRPQALDCLLAGCESTITLPDVAQOMLNPVESHIEKFEHDMMAFGTT
           HL"
promoter   complement(4922..4949)
           /note="promoter matrix score of 47; putative; within ORF
           f220, which would suggest alternate start codon"
           complement(4962..7097)
CDS       /note="similar to phosphotransferase system enzyme I"
           /codon_start=1
           /transl_table=11
           /label_ORF_f711
           /protein_id="AAC43053.1"
           /db_xref="GI:409787"
           /translation="WCSGSAGGILTPISLIDNALGNLPAKGVDAEQSALENGLTV
           LKNIERLSDSGATSALIEAHRSLAGDTSLEBHLAAGSAGISCEALIVASANHCE
           EFSRSSSYLQSRALDVRVCFOLQOITGEORFPAPGKLTQPAICMADELTPSOELE
           LDKNLKGLLSGGTSHVILARSFNIPLVGVVIDALTPMOOQITIIDGNAGATV
           VERGEAVARYOQEARVODALREOQVWLTQOARPDRIETIAINIAHSVEQAAG
           NGAEGVGLFRTMLYMDRTSAGESELYNIFCOALESANGRSIYATNMDIGCKPVDY
           LNIPEANPFLIGRAVRIEYASLETTQLRSILRASAGSLKIMIPMISMEELIW
           KEKLAQAQOLRNENHLPDEKIQIGIMLEPVSVMIIDCCCEIDFESIGSNDLQYL
           LAVDRNNAKVTNRHNSLNPALFDYAAQVAVHROGKIGICELGAKGSVLPILVGL
           GLDELMSAPSIIPAKARMAQDRECRKLNOAMACRPSLEVEHLAOFRTMOODAP
           LYTAECTILESDMRSKEVILKGMTDLILAGRCRYRKLLEADILAREAVFSTGICPSF
           AIPHKSSEHIEOSTISVARLOAPVRMGDEAOPIIMLTINKHAQGHNRITSRILAR
           IMHEERNALVNAASADAIALSLDHELEL"
CDS       7099..7428
           /codon_start=1
           /transl_table=11
           /label_ORF_o109
           /protein_id="AAC43054.1"
           /db_xref="GI:409788"
           /translation="MRARIICGFRIVSETGSSSSSLLTSASAGASGMSRSRHLKR
           WCASCSASPEISNMQLSPASVSPISARALLALPCQCAGATQFVWKRYTAFOGL
           PDAHYRR"
promoter   complement(7108..7136)
           /note="promoter matrix score of 46; putative"
misc_feature
           7200..72839
           /note="corresponds to lambda clone EC21-52"
terminator 7578..7699
           /note="putative"
```

```
promoter   7771..7800
           /note="promoter matrix score of 48; putative; within ORF
           0359, which would suggest alternate start codon"
           7771..8850
CDS       /note="similar to phosphotransferase system enzyme II"
           /codon_start=1
           /transl_table=11
           /label_ORF_o359
           /protein_id="AAC43055.1"
           /db_xref="GI:396296"
           /translation="MNEIVQILKTRQHLMTGVSHNIPFVSGGILLAVSYMLYKGA
           VPDAAVDPMILKRLFDIGVAGLTMVPLFAAYIGYSIAERSALAPALGAWGNSFCAG

Query Match      27.9%; Score 500; DB 1; Length 176195;
Best Local Similarity 97.1%; Pred. No. 1.9e-106;
Matches 509; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 1267 agagtcgacctgcagcagcatgcaagcttg99ltccaccctgaccccatgcccgaactcagaag 1326
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36873 AGAATTTGGCTGGCGGAGTAGAGCGGGGTGCTCCACCTGACCCCATGCCAGACTCAGAAAG 36932

Qy 1327 tgaagcgccgtagcgcgcagatgtagtgggtggtcctcccatgcgagatgaggaactgcc 1386
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36933 TGAACGCCCTACCGCGGATGTAAGTGTGGGTCTCCCATGCGAGTAGGGAACCTGCC 36992

Qy 1387 aggcattcaataaaacgaagagctcagtcgaagaagctggcctttcttatctgttt 1446
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36993 AGGCATCAATAAATAACGAAGAGCTCAGTCGAAGAGCTGGGCTTTGCTTTATCTGTGTG 37052

Qy 1447 ttgctggtgaagcgtctcctcctgaatagaacaatcgcgcggaagcgaattgaacgttgcg 1506
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37053 TTGTCGGTGAAGCCTCTCCGAGTAGAGCAAAATCCGCGGAGCGGATTTGAACGTTGCG 37112

Qy 1507. aagcaacgcccggagggtgcgggcgaagacgcccgcataaacctgcaggaatcaaat 1566
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37113 AACCAACGCGCCCGAGGAGGTGGCGGCGAGCGCCGCATTAACCTGCCAGCATCAAAAT 37172

Qy 1567 aagcagaagccatccctcgaagagatggccttttggcgtttctaaactctcttcgctgc 1626
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37173 AACCGAAGGCCATCTCGACGAGATGGCTTTTGGCTTCTTCAAACTCTTCTGTCGTC 37232

Qy 1627 atactcaagcattcccccacagatacgaatgaactagctcgttttgcacagaaga 1686
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37233 ATATCTTCAAGCCATCCGCCACAGATAGCGTAACTAGCTGTTTTCATCAGAGAA 37292

Qy 1687 gcaagctatgaacacactcttaaaacccctggaacacattggcatatgaatcagtgccag 1746
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37293 GCAGCTATGAACCACTCTTAAAAACCTGGAAACATTTGGCATTTGATCAATGCTCAG 37352

Qy 1747 caaatgtagtggtccgaagcgaacaacataactactcaatgcccg 1790
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37353 CAACTGTATGTGCTCGAAGCAACCAATTAATCTCAATGCTG 37396

RESULT 10
SYNCGNABP LOCUS SYNCGNABP 860 bp DNA linear SYN 27-APR-1993
DEFINITION e.coli rna operon promoters fused to rnb operon terminators.
ACCESSION K00764
VERSION K00764.1 GI:209262
KEYWORDS 16S ribosomal RNA; 5S ribosomal RNA; promoter region; ribosomal
SOURCE RNA: rna operon; rnb operon; terminator.
ORGANISM escherichia coli dna; clones psl and p55.
REFERENCE 1 (bases 1 to 860)
AUTHORS Sarmientos,P., Sylvester,J.E., Contente,S. and Cashel,M.
TITLE differential stringent control of the tandem e. coli ribosomal rna
promoters from the rna operon expressed in vivo in multicopy
plasmids
JOURNAL Cell 32, 1337-1346 (1983)
MEDLINE 83180429
COMMENT [1] assessed stringent control of the p1 and p2 promoted
```


terminator	2039. .2066	/standard_name="rnb r2"	/citation=(2)	complement(3267)	/note="base 3267 represents the first base of the newly synthesized strand"
rep_origin					
gene				complement(4026. .4956)	/direction=left
CDS				/gene="bla"	complement(4026. .4886)
				/gene="bla"	/codon_start=-1
				/transl_table=1	/product="beta-lactamase"
				/protein_id="AAC53597.1"	/db_xref="GI:595747"
					/translation="MSIQHFRVALIIPFPAFLPVFHPETLVKVKAEQDLGARY
					IEDLSNGKILIESFRPEERPEPMSTFVLLCGAVLSHVDAGDQLGRIHYSNDLVE
					YSPVTEKHLLDGMVTRELCSAATMSMTANLLITIGPKELTAFLAHMGDHYRL
					DYRPELNEAIPNDERDTPPAAATTLRLILGELLTLASROOLIMMADKVAQPL
					LRSLPAGMFWIADKSGAGERSRGIIALPDGKPSRIIVITYTGSQATMDERNRQIA
					ELASLIKHW"
promoter				complement(4928. .4956)	
BASE COUNT	1293 a	1290 c	1290 g	1221 t	
ORIGIN					
Query Match	22.5%	Score 404;	DB 12;	Length 5094;	
Best Local Similarity	96.5%	Pred. No. 5e-84;			
Matches 413;	Conservative	0;	Mismatches 15;	Indels 0;	Gaps 0;
OY	1267	agaagtcgacccctggagagcctgcaagctgttggtcccaaccctgcagcccatgcggaactcagaag	1326		
Db	1237	AGATTTCCTCTGGCGCGCGCTAGCGCGGTGGTCCACCTGACCCCAATGCCGAACCTCAGAAg	1296		
OY	1327	tgaaacgcccgtagcgcgaatgtatgttggggtctccccaatgcgagagtaaggaaactgcc	1386		
Db	1297	TGAAACGCCCGTAGCGCCGATGTAAGTGTGGGGTCTCCCATGCGAGATAGGGAACGTGCC	1356		
OY	1387	agggcatcaataaaaaaagaaaggtctcagtcgaaagaactctggcccttcgtttatctgtgtc	1446		
Db	1357	AGGCATCAAAATAAACGAAGAGGCTCAGTCGAAGAACTGGGCTTTCGTTTATCTGTTGT	1416		
OY	1447	ttgtcgtgtgaacgcgtctctctcctagtatgaggaacaatccgcgcgggagcgagatttgaacgttcg	1506		
Db	1417	TTTGCGTGAAACGCTCTCTCTGTAAGTAGCAAAATCCGCGGGAGCGGATTTGAACTTCGCG	1476		
OY	1507	aagcaacgcccgcgagaggttgcgcggaagaagcccgccataaactctgcagagcatcaaat	1566		
Db	1477	AAGCAAGCGCCCGAGGGGTGGCGGGCAGAGCGCCGCATAACTGCGCAGCATCAAAATT	1536		
OY	1567	aagcagaagcccatcctctgaagcgatgagccttcttcgtttctaaaaaactcttcctgtgcgc	1626		
Db	1537	AAGCGAAGGCGCATCTCAGCAGATGGCCCTTTTGGCTTCAACAAACTTCTGTCGTC	1596		
OY	1627	atatctaaagcatccccccacagataacggttaactagctcgtttttgatcatgagaaa	1686		
Db	1597	ATATCTACAGGCGATCCGCCACAGATACGGTAACACTAGCCTCGTTTTGCAATCAGGAAA	1656		
OY	1687	gcagcctat	1694		
Db	1657	GCAGCTGT	1664		
RESULT 12	AF216802/c	6733 bp	DNA	Linear	SYN 22-JUN-2000
LOCUS	AF216802				
DEFINITION	Shuttle vector pDL278, complete sequence.				
ACCESSION	AF216802				
VERSION	AF216802.1				
KEYWORDS	Shuttle vector pDL278.				
SOURCE					

ORGANISM	Shuttle vector pDL278
REFERENCE	artificial sequence: vectors.
AUTHORS	1 (bases 1 to 6733)
TITLE	LeBlanc,D.J., Lee,L.N. and Abu-Al-Jaibat,A. Molecular, genetic, and functional analysis of the basic replicon of pVA380-1, a plasmid of oral streptococcal origin
JOURNAL	Plasmid 28 (2), 130-145 (1992)
MEDLINE	93028887
PUBMED	1409970
REFERENCE	2 (bases 1 to 6733)
AUTHORS	Dunny,G.M. and Bae,T.
TITLE	Direct Submission
JOURNAL	Submitted (15-DEC-1999) Microbiology, University of Minnesota, Box 196 UMMC, 1460 Mayo Building, 420 Delaware St. SE, Minneapolis, MN 55455, USA
FEATURES	Location/Qualifiers
source	1..6733 /organism="Shuttle vector pDL278" /db_xref="taxon:111473" /focus /note="Gram positive-Gram negative shuttle vector"
CDS	1682..2449 /codon_start=1 /transl_table=1 /product="spectinomycin adenyltransferase" /protein_id="AAE78194.1" /db_xref="GI:8650403" /translation="MRRIYNTYTEQIRKKVKILRKHLKNNLIGTYMGSGVESGLRPN SDLDLVVSEPIIDOSKEITLQIRPSIKIGDKSNLRILETLIIQDMVPNNHP KOEIFXGWLQELDEOGRIPQKELNSDUTIMLYQAKRNKRIRIGNDLELPDPFS DVRRAIWDSSEELIDNYODETNSILTCMLGMATMDGKIIPRDIAGNAAVESPLEH RERLIARVSYLEGIENWTENNVLNLTINYLNINRLKL" complement(4639..4878) /note="ORF 1" /codon_start=1 /transl_table=1 /product="unknown protein" /protein_id="AAF78195.1" /db_xref="GI:8650403" /translation="MSKEELALTIPTPKGVFGKEYVELTKDFEFEGFLIYSNRNLVHQ KLENEOLRIVPLRKSRKRPASWNELKKVRERA" complement(5042..6183) /standard_name="rrnB r172 terminator" complement(5042..6422) /organism="Escherichia coli" /db_xref="taxon:562" complement(<6183..6422) /codon_start=1 /transl_table=1 /product="beta-galactosidase alpha peptide" /protein_id="AAF78196.1" /db_xref="GI:8650404" /translation="MTMTSPSLAACRSTLEDPRYPSSNSLAIVYLQRDMENPGVTQLN RLAHPFPASWRNSEEARTRPSSQLRSLSNGECREL"
BASE COUNT	1926 a 1478 c 1345 g 1984 t
ORIGIN	
Query Match	22.5%; Score 404; DB 12; Length 6733;
Best local similarity	96.5%; Pred. No. 5e-84;
Matches 413; Conservative %	0; Mismatches 15; Indels 0; Gaps 0;
Db	5984 AGAATTCTCCCTGGCGCGAGTAGCGGCGGTGCCACCTGCACCCTCAGCAACTCAGAAG 5925
Oy	1327 TGAAGACGCGCTAGCGCCATGTAAGTGTGGGGGTCTCCCATGCGAGAAGTACGGAACCTGCC 5865
Db	5924 TGAAGACGCGCTAGCGCCATGTAAGTGTGGGGGTCTCCCATGCGAGAAGTACGGAACCTGCC 5865
Oy	1387 aggcatacaataaaagcgaaggtctcagtcgaaagaactggcgcttgcgtttatatcgttgt 1446
Db	5864 AGGCATCAAAATAAAGCAAGAGGCTCAGCGCAAGAGCTGGCGCTTGCGTTTATCTGTTGT 5805

QY 1447 ttgtcgtgtaacgctctcctcgtagtaggacaacatccgcggagcgagatttgaaactgtcg 1506
|||||
Db 5804 TTGTCGCTGAACGCTCTCTGAGTAGAGCAAAATCCGCGGAGCGGATTGGAACGTGGCG 5745
QY 1507 aagcaaacgcccggagggtggcgagcgccgcacataaactgccaagcaacaatt 1566
|||||
Db 5744 AAGCAACGGCCCGAGGCTGGCGGAGAGCGCCGCCATTAACCTGCGAGCATCAAAATP 5685
QY 1567 aagcaagaagccatccctgaagatgaccttttgcgtttctacaacactctctcgtcgc 1626
|||||
Db 5684 AAGCAGAAAGCCATCTCTGAGAGATGGCCCTTTTGGCTTCTACAAACTCTTCTGCTGTC 5625
QY 1627 atatctcaagccatcccccacagaatacgtaaacatcgcttttgcatacaggaaa 1686
|||||
Db 5624 ATATCTCAAGCCATCCGCCACAGATACGTAACCTAGCCTCTTTTGCATCAGAGAAA 5565
QY 1687 gcaagctat 1694
|||||
Db 5564 GCAAGCTGT 5557

RESULT 13
AF216803 6984 bp DNA circular SYN 01-DEC-2000
LOCUS AF216803
DEFINITION Shuttle vector pDL276, complete sequence.
ACCESSION AF216803
VERSION AF216803.1 GI:10179615
KEYWORDS Shuttle vector pDL276.
SOURCE Shuttle vector pDL276.
ORGANISM artificial sequence; vectors.
REFERENCE 1 (bases 1 to 6984)
AUTHORS Dunny,G.M., Lee,I.N. and Leblanc,D.J.
TITLE Improved electroporation and cloning vector system for
gram-positive bacteria
JOURNAL Appl. Environ. Microbiol. 57 (4), 1194-1201 (1991)
MEDLINE 91282471
PUBMED 1903518
REFERENCE 2 (bases 1 to 6984)
AUTHORS Dunny,G.M. and Bae,T.
TITLE Direct Submission
JOURNAL Submitted (15-DEC-1999) Microbiology, University of Minnesota, Box
196 UMC, 1460 Mayo Building, 420 Delaware St. SE, Minneapolis, MN
55455, USA

FEATURES
source
1. 6984
Location/Qualifiers
/organism="Shuttle vector pDL276"
/db_xref="taxon:111474"
/note="Gram positive-Gram negative shuttle vector"
1746..2540
/codon_start=1
/transl_table=11
/product="aminoglycoside transferase"
/protein_id="AAG13816.1"
/db_xref="GI:10179616"
/translation="MAKKRISPELKKLIKRYCVDTEGSMSPAKYKLVGENENTLYK
MTDSRYKGTIVDEREKDMLLEGLPVPKYLHFRHGWSNLMSNDADVLCSEY
EDEOSPEKIIELVAEICIRLFHSIDISDPTNSLRHAEILYLLNNDLADVCEWME
EDTPPEKPRELDELKTEKEPEELVSHGDLGNSNFFVKGVSFIDIGRGRARAKW
VDIATCVSRIRREDIGEQYVELFPDLGIRKPRMEKIKYIILDELFL"

BASE COUNT 1869 a 1593 c 1510 g 2012 t
ORIGIN

Query Match 22.5%: Score 404: DB 12: Length 6984:
Best Local Similarity 96.5%: Pred. No. 5e-84:
Matches 413: Conservative 0: Mismatches 15: Indels 0: Gaps 0:

QY 1327 tgaacgcgtaagcccgatgtagtgggtctcccatcgcagagtagggaactgcc 1386
|||||
Db 6175 TGAACGCCCGTAGCGCGGATGAGTGTGGGGTCTCCCATGCAAGATGAGGAATCGCC 6116
QY 1387 aggcataaataaagaaagcccaagtcgaagagacaggcccttcgtttatctgtgtc 1446
|||||
Db 6115 AGGCATCAAAATAAACAAGGCTCACTGCAAAACATGGGCTCTTTTATCTGTTGT 6056
QY 1447 ttgtcgtgtaacgctctcctcgtagtaggacaacatccgcggagcgagatttgaaactgtcg 1506
|||||
Db 6055 TTGTCGCTGAACGCTCTCTGAGTAGAGCAAAATCCGCGGAGCGGATTGGAACGTGGCG 5996
QY 1507 aagcaaacgcccggagggtggcgagcgccgcacataaactgccaagcaacaatt 1566
|||||
Db 5995 AAGCAACGGCCCGAGGCTGGCGGAGAGCGCCGCCATTAACCTGCGAGCATCAAAATP 5936
QY 1567 aagcaagaagccatccctgaagatgaccttttgcgtttctacaacactctctcgtcgc 1626
|||||
Db 5935 AAGCAGAAAGCCATCTCTGAGAGATGGCCCTTTTGGCTTCTACAAACTCTTCTGCTGTC 5876
QY 1627 atatctcaagccatcccccacagaatacgtaaacatcgcttttgcatacaggaaa 1686
|||||
Db 5875 ATATCTCAAGCCATCCGCCACAGATACGTAACCTAGCCTCTTTTGCATCAGAGAAA 5816
QY 1687 gcaagctat 1694
|||||
Db 5815 GCAAGCTGT 5808

RESULT 14
SYNPTL61T 8565 bp DNA circular SYN 24-MAY-1994
LOCUS SYNPTL61T
DEFINITION Cloning vector pTL61T, complete sequence.
ACCESSION M29896
VERSION M29896.1 GI:209190
KEYWORDS cloning vector; synthetic sequence.
SOURCE Synthetic plasmid pTL61T (natural host E. coli).
ORGANISM unidentified cloning vector.
REFERENCE 1 (bases 1 to 8565)
AUTHORS Lim,T. and Pierre,R.
TITLE Improved vector system for constructing transcriptional fusions
that ensures independent translation of lacZ
JOURNAL J. Bacteriol. 172, 1077-1084 (1990)
MEDLINE 90130263
COMMENT Draft entry and computer-readable sequence for [1] kindly submitted
by T.G. Lim, 15-NOV-1989.
Authors indicate an uncharacterized mutation between bp 3274-3379,
which alters a EcoRI site normally present in the E. coli lacZ gene.

FEATURES
source
1. 8565
Location/Qualifiers
/organism="unidentified cloning vector"
/sub_species="Cloning vector pTL61T"
/db_xref="taxon:45196"
1..92
/note="multiple cloning site"
116..192
/note="RnaeIII processing site"
238..239
/organism="Cloning vector pTL61T"
256..3330
/note="beta-d-galactosidase"
/codon_start=1
/transl_table=11
/protein_id="AA72803.1"
/db_xref="GI:209191"
/translation="MTMTDSLAAYVLIQRDMENPGYTOINRLAAHPPPASRNSEAR
TDRPSOQLSLNGEMRFAPPAPEAVPESWLECDLPEADTVVVPSPNQWHGYDAPLYT
NVTYPTVNPVPVPTENPGCYSLTFNVNDESLQSGORFIIFDVNAAFFHLNCSGRNV
GYGDSRLPSEFDSLAFLRAGENRLAAVYLIRSDSYLEDODMAMSGIFRVSLLHK
PTTQISDFHATRFNDRAVLEAVQWCGELRDYLRVTVSLMOGEGOVAAGTAPFG
GRTIDBERGYADRVTLRNVENPKLMSAETPILYAAVYELHATADCTLEAFCQVCFR
EVRIENGLLLNGKPLILRGVNRHEHPLHGVQWDEQTKVQDILILKQNNPNAVPCSH

misc_recomb	3392..3393	/organism="Cloning vector pTL61t"
misc_recomb	3890..3891	/organism="Cloning vector pTL61t"
misc_recomb	3971..4470	/organism="Cloning vector pTL61t"
misc_recomb	4090..4588	/product="5S ribosomal RNA"
misc_recomb	4866..4887	/product="5S ribosomal RNA"
variation	7071..7073	/organism="Cloning vector pTL61t"
variation	7824..7826	/note="tgc in pBR322; tac in pTL61t"
misc_recomb	7826..7827	/note="a in pBR322; att in pTL61t"
misc_recomb	join(8565..1)	/organism="Cloning vector pTL61t"
misc_recomb	/note="pRS51 DNA end/synthetic DNA start"	
misc_recomb	/organism="Cloning vector pTL61t"	
BASE COUNT	2028 a 2266 c 2333 g 1938 t	
ORIGIN		

Query Match	22.5%	Score 404:	DB 12:	Length 8555:
Best Local Similarity	96.5%	Pred. No. 5.1e 84:		
Matches 413: Conservative	0:	Mismatches 15:	Indels 0:	Gaps 0:
Qy 1267 agagtcgacgcctgcacgacatgcgaagcttgggtccaccctgcaccccatgcgcgaactcagaag				
Db 3965 AGAATTTCCTCGGCGCGAGTACGGCGGGTGTCTCCACMCTGACCCCATGCCAGACTCAGAAAG				
Qy 1327 tgaacgcgcgttagcgcgatgtagtgttgggtgtctcccatgcgaagtaaggaaactgcc				
Db 4025 TGAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCATCGCAGAGTAGGGAAC TGCC				
Qy 1387 aggcatacaataaaacgaaagcttaagtcgcgaagactggccttctgtttatctcgttgt				
Db 4085 AGGCTCAAAATPAAACGAAAGGCTAGTCGAAAGACTGGGCCCTTTCGTTTATCTGTGT				
Qy 1447 ttgtcgtgtgaacgcctctcctgaataggaacaaatccgcgcggagcgagattgaacgttgcg				
Db 4145 TTGTGGGTGAACGCTCTCTCTGAGTAGAGACAATTCGCGCGGAGCGGATTTGAACTGTGCG				
Qy 1507 aagcaacgcccgcgggggttggcgggcgcaggaacgcccgcacataaactgcacgacatcaalt				
Db 4205 AAGCAACGCCCGGAGGGGTGGCGGCGCAGAGACCCGCCCATAAATGCGAGCATCAAAATT				
Qy 1567 aagcagaagagccatctctgcagatgtagcttlttgcgtlttcaaaactctctctgcgc				
Db 4265 AAGCGAAGGGCATCTCTGACGAGATGGCCCTTTTGGGTTCTTCAAAACTCTTCTCTGCTGC				
Qy 1627 atattcaaaccatccccccacagaatcgcgaaactgcgcgttttgcatacgaagaa				
Db 4325 ATATTCAACGCATCCCCCACAGATACGGTAACCTAGCCTGATTTTGCATCAGGANA				
Qy 1687 gcaagctat 1694				
Db 4385 GCAGCTGT 4392				

LOCUS	CYPCG1408	9113 bp	DNA	circular	SYN 01-DEC-2000
DEFINITION	Promoter-probe vector pcg1408, complete sequence.				
ACCESSION	U01228				
VERSION	U01228.1	GI:885956			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
1	Cloning vector PCG1408.				
2	Cloning vector PCG1408				
3	artificial sequence: vectors.				
4	1 (bases 1 to 9113)				
5	David,M., Bartz,S.U., Weaver,J.D. and Kelley,J.L.				
6	Isolation and Characterization of Promoter Sequences from				
7	Clavibacter xyli subsp. cynodontis				
8	Unpublished				
9	2 (sites)				
10	Taylor,J., Stearman,R.S. and Urrutani,B.B.				
11	Development of a native plasmid as a cloning vector in Clavibacter				
12	xyli subsp. cynodontis				
13	Plasmid 29 (3), 241-244 (1993)				
14	93361581				
15	3 (bases 1 to 9113)				
16	David,M.				
17	Direct Submission				
18	Submitted (17-FEB-1995) Michael David, Crop Genetics, 10150 Old				
19	Columbia Rd, Columbia, MD 21046, USA, 21046				
20	Location/Qualifiers				
21	1..9113				

```

CDS
    88..747
    /codon_start=1
    /transl_table=11
    /product="chloramphenicol acetyltransferase"
    /protein_id="AA070410.1"
    /db_xref="GI:885957"
    /translation="MEKKITGYTTVDISQSHRKEFEAFQSVAACTYNOTVOLDITAF
    LKTYKKNKHEFPATHTILARLMANHPPEPRAMKDGELVWDSVHPCYTVHEQTEFFEF
    SLTMEYHDDPEFLHIYSOVACGENLAYEPKGTIENMFVSANPWSTSDLVNW
    ANMDQFAPVPTVMGKYTVQGDVKVLPPLAIQVHHAVCDSPFHGRRLNELQYCDBMQGG
    A"
terminator
    1185..1228
    /note="T1"
terminator
    1360..1388
    /note="T2"
terminator
    1684..1726
    /note="T1"
terminator
    1858..1885
    /note="T2"
CDS
    2702..3562
    /codon_start=1
    /transl_table=11
    /product="beta-lactamase"
    /protein_id="AA070411.1"
    /db_xref="GI:885958"
    /translation="MSIOHFRVALIPPEAPACIPAFAPETLVKVAEDQGLDYVE
    IELDNGSKILSEFRPEERFPMASTFKVLLGAYLSRYDAQEQLGRITIHQSNDVRE
    TSPVLEKHLDGDMTVRELCSAAITWSDMTAANLLTTIGGKRELTAFLHNHGVTRT
    RWSELPEDMEALPNDERDPTMPAAMATTLRLKLTGELTLASRQQLIDMEEDKVAAGPI
    LRSALPACWFLFIADKSGAGERGSRGIIALIGDPKSRIVIVITTTGSQATMERNRROI
    AEGASLIKHH"
    4084..4508
    /note="E. coli origin of replication"
rep-origin
    5465..8413
    /note="Claybacter origin of replication"
    /citation=11
    /citation=12
    /citation=12
    join(8414..9113,1..5464)
    /organism="Cloning vector pKK232-8"
    /note="region derived from vector pKK232-8, GenBank
    accession Number U13859"
    /db_xref="taxon:74582"
    8947..8990
    /note="T1"
terminator

```

```
misc_feature      join(9109..9113,1..26)
                   /note="polylinker"
BASE COUNT        2113 a      2644 c      2396 g      1960 t
ORIGIN
```